Probabilistic Characterization of Structural Demand Under Earthquake Loading – Revisiting the Lognormality Assumption

X. Romão & R. Delgado
Faculty of Engineering of the University of Porto, Portugal

A. Costa
University of Aveiro, Portugal

ABSTRACT:
Probabilistic methods are widely used to assess the performance of structures under earthquake loading and require an adequate probabilistic characterization of the structural demand. Currently, many existing seismic safety assessment studies are developed under the assumption that structural demand conditional to a given seismic intensity follows a lognormal distribution. Given the importance of this assumption, an in-depth analysis of its validity is carried out for several case studies using adequate statistical methods. This assessment of the probabilistic demand distribution type is based on the analysis of 5 reinforced concrete framed structures subjected to sets of 50 real ground motion records scaled to several intensities. The structural demand parameters addressed by the study are the chord rotation, the curvature, the shear force and the inter-storey drift.

Keywords: probabilistic seismic demand, normality testing, reinforced concrete structures, seismic analysis

1. INTRODUCTION

In the framework of Performance Based Earthquake Engineering and of the development of methodologies for seismic risk reduction, probabilistic methods are seen as superior means of assessing the performance of structures under earthquake loading. In this context, a common assumption is that, for some level of the earthquake intensity measure (IM), the probability distribution of a structural demand parameter can be modelled by a lognormal distribution. Although this hypothesis is found in numerous research studies (e.g. see Shome and Cornell 1999, Cornell *et al.* 2000, Goda *et al.* 2009 Ruiz-Garcia and Miranda 2010), its consistent assessment using adequate statistical tools has not been carried out yet. Therefore, a study addressing the probabilistic distribution of several demand parameters is proposed herein to evaluate the hypothesis that the referred probabilistic demand could be adequately modelled by a lognormal distribution. In addition, the suitability of the normal distribution is also assessed.

2. DESCRIPTION OF THE PROPOSED STUDY

The presented study addresses the probabilistic demand distributions obtained from the analysis of 5 reinforced concrete (RC) structures subjected to earthquake records of increasing intensities. The selected demand parameters are the maxima of the section curvature φ, chord rotation θ and shear force V, and of the inter-storey drift over the height of the structure Δ. The selected structures were analysed for suites of 50 ground motions scaled for several intensities in order to evaluate the referred hypotheses for different hazard levels. Furthermore, 2 different IMs - the peak ground acceleration (PGA) and the 5% damping spectral acceleration of the ground motion for the fundamental period of the structure T_1 ($S_a(T_1)$) - were also considered to evaluate the influence of this parameter. Statistical tests were applied to the demand samples to evaluate the adequacy of the distribution hypothesis. It is noted that the tests only aim to determine the validity of the hypothesis regarding the type of probabilistic distribution and do not make any inference about their parameters.

The considered tests were selected according to the results of a benchmark efficiency test previously carried out (Romão *et al.* 2010) and were divided into 2 groups. The first group has tests suited to identify non-normal distributions when the data is symmetric and the second group has tests suited to deal with normally distributed data that has outliers. The tests of the first group are the β_3^2 test (Coin 2008), based on a polynomial regression, the R_{sJ} test (Gel *et al.* 2007), based on the ratio of the standard deviation and a robust measure of dispersion, and the T_w test (Bonett and Seier 2002), based on a modified measure of kurtosis. The tests of the second group are the *CS* test (Chen and Shapiro 1995), based on normalized spacings, the $T_{TLmom}^{(t)}$ test (Romão *et al.* 2010), based on the robust generalization of the sample L-moments defined as the trimmed L-moments (Elamir and Seheult 2003), and the $T_{MC\text{-}LR}$ test (Brys *et al.* 2008) based on robust measures of skewness and tailweight.

The 5 RC structures that were analysed are the four-storey, three-bay ICONS frame (Carvalho *et al.* 1999), the two six-storey frames presented in (Ferracuti *et al.* 2009), and two ten-storey frames presented in (Athanassiadou 2008) named FRH and FRH-2 - the regular and the irregular frames are referred herein as REG10 and IRREG10, respectively. Figure 1 presents the elevation views of the selected structures. Details about the frame characteristics are found in the previously cited references.

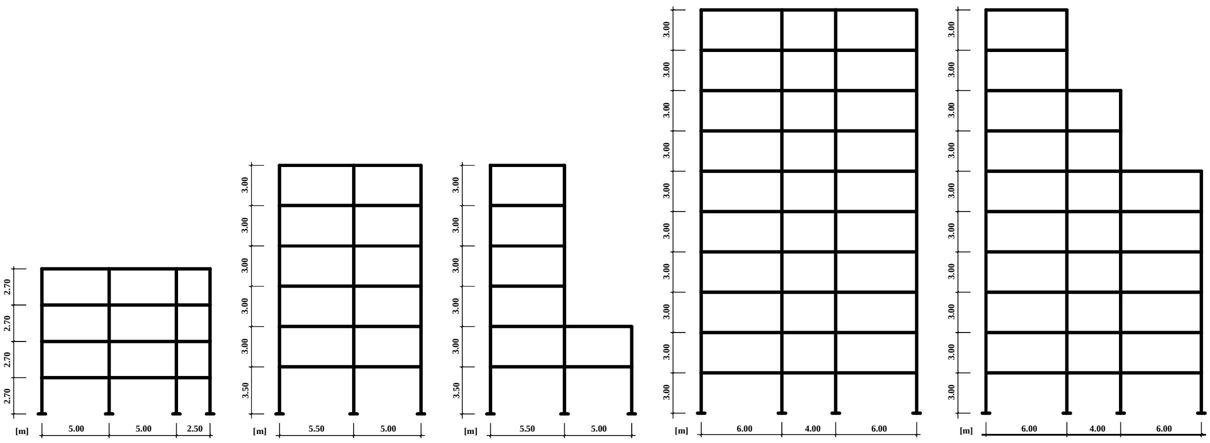

Figure 1. Elevation views of the considered structures.

The nonlinear response analysis of the frames under earthquake loading was carried out using a two-dimensional analysis platform (Varum 1997). The considered programme is able to model the behaviour nonlinearities of beams and columns, as well as the large lateral deformations which are simulated by a leaning column that reproduces the destabilizing P-Δ effects of the gravity loading. Elements are modelled with plastic hinges located at the member ends, where inelastic flexural behaviour is considered. The inelastic behaviour of the plastic hinges is defined by moment-curvature relations (Arêde and Pinto 1996) based on mean material property values. Hysteretic behaviour of the members was modelled by the piecewise linear Costa-Costa model (CEB 1996), considering stiffness degradation and pinching effects. The plastic hinge length values were considered equal to the member cross section depth for beams and equal to half of the member cross section depth for columns. Viscous damping was assumed to be proportional to initial stiffness with a parameter calculated for the first mode period of the structures and for 2% of the critical damping.

The seismic demand considered for each structure consisted of a suite of 50 real ground motions extracted from the Pacific Earthquake Engineering Research Center NGA database (NGA 2009). Each structure was analysed for the selected ground motions scaled for 9 intensities in order to reflect different return periods (RPs). The selected RPs were 37, 73, 95, 225, 475, 976, 1980, 2480 and 4950 years. The reference seismic scenario selected to define the scaling factors for each RP corresponds to that of Zone 3 of the Portuguese territory, considering the intraplate seismic action and a soil of type B according to the Portuguese National Annex of Eurocode 8 (EC8-1 2009). The PGA considered for this scenario was 0.17*g*, corresponding to a RP of 475 years. The PGA values associated to the other RPs were obtained based on the results of the hazard studies presented in (Carvalho *et al.* 2008). The selection of the ground motion records was based on several criteria aiming to minimize the positive and negative (absolute values) mismatches between the response spectrum of the real record, scaled by a scaling factor between 0.75 and 1.33, and the reference response spectrum, over the period range between T_2 and $1.5T_1$, where T_2 is the period of the second mode of the structure under consideration. After selecting the 50 ground motions for a given structure, these were then scaled for the values of the selected IMs matching the RPs previously referred.

3. RESULTS OF THE GOODNESS-OF-FIT ASSESSMENT

General conclusions regarding the assessment of the selected statistical distribution hypotheses are presented in the following. For the sake of brevity, only a sample of the results is shown, along with representative figures illustrating the more important findings. The goodness-of-fit (GOF) results are presented separately for the beam and column demand data and for the previously referred demand parameters (φ, θ, V and Δ). For the φ and θ demand, tests were applied separately for positive and negative data and the presented results are the average of the results obtained for both signs.

Results are presented in terms of average percentage of acceptance (APA) data for the considered levels of seismic intensity and for both the normal and the lognormal distribution hypotheses. The APA represents the number of times a certain group of tests does not reject a given distribution hypothesis, considering a confidence level of 95%. In quantitative terms, a 75% threshold APA was considered as the limit value above which a given distribution hypothesis is accepted to be appropriate to model the probabilistic distribution of a parameter. Since demand distributions are not expected to follow a theoretical statistical model perfectly, such limit is considered to be adequate to represent the average contribution of the control sections of all the structures. With respect to the selected groups of tests, the following three groups are defined: Group 1 – Tests for symmetric data; Group 2 – Tests for data with potential outliers; Group 3 – All the tests from Group 1 and from Group 2

The GOF results obtained from the application of the tests from Group 3 to the column and beam demand datasets of all the structures are presented in Figs 2 a) and b), respectively, for the case where PGA is the IM. These results show that the lognormal distribution hypothesis yields better results for the probabilistic modelling of the φ and θ demand distributions. For the probabilistic representation of the V demand, both distribution hypotheses yield similar results. However, the APA results of some demand parameters do not meet the 75% threshold for several seismic intensities, particularly for the V demand in beams. Figs 3 a) and b) present results similar to those of Figs 2 a) and b) now for the case where $S_a(T_1)$ is the IM. As for the previous case, the lognormal distribution is more adequate to model the probabilistic distribution of the φ and θ demand. With respect to the V demand, again both distribution hypotheses yield similar results. More importantly, the observation of these results allows concluding that $S_a(T_1)$ leads to higher APA results. Moreover, for the φ and θ demands, such APA results meet the 75% threshold for most seismic intensities. However, for the case of V demand, there are APA values below the referred threshold for several seismic intensities, particularly in beams. With respect to the Δ demand, Figs 4 a) and b) present the GOF results obtained from the application of the tests from Group 3 to the datasets of all the structures for the cases where PGA and $S_a(T_1)$ are the IM, respectively. The presented results indicate that the lognormal distribution is generally more adequate than the normal distribution to represent the probabilistic distribution of the Δ demand. Furthermore, the advantage of one IM over the other is not as clear as for the previous demand parameters. Still, $S_a(T_1)$ is favoured since it leads to higher APA results for the higher seismic intensity levels. Nonetheless, there are some APA values below the 75% threshold.

Globally, the results indicate that the lognormal distribution could be suitable for the probabilistic modelling of the φ and θ demand of beams and columns, as well as for the probabilistic modelling of the Δ demand. With respect to the V demand, the results indicate that both the normal and the lognormal distributions may have the same potential to model the probabilistic distribution of this parameter. Moreover, $S_a(T_1)$ is seen to be a more adequate IM for the purpose of obtaining demand distributions more compatible with the referred distribution hypotheses. However, the GOF results are not totally satisfactory since there are APA values below the 75% limit for some seismic intensities.

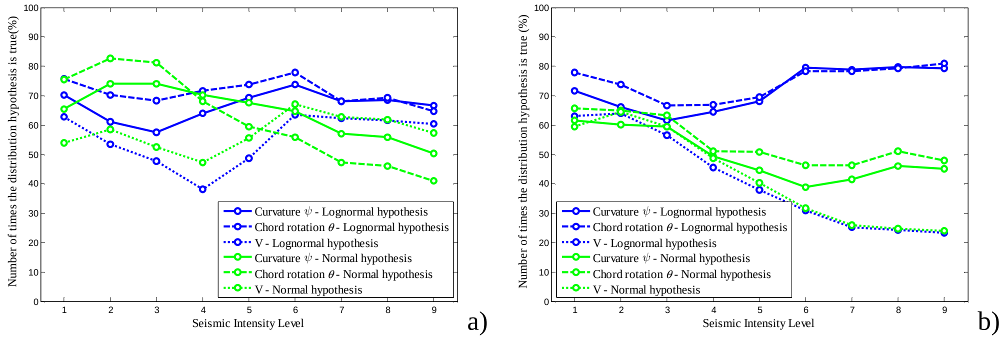

Figure 2. APA results from the tests of Group 3 for the column (a) and beam (b) datasets when PGA is the IM.

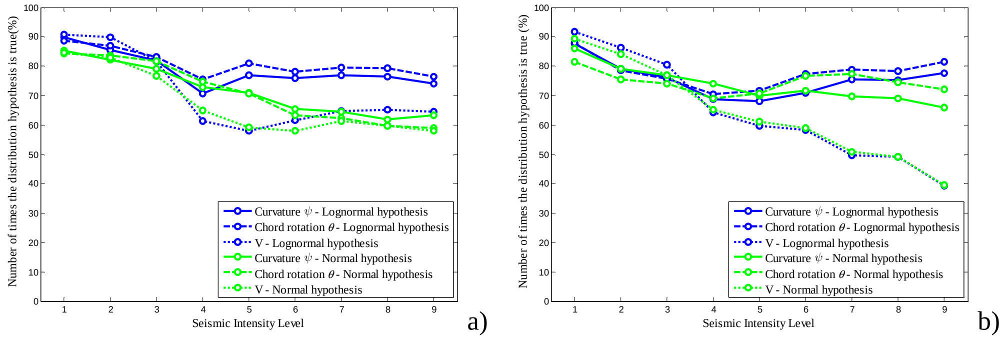

Figure 3. APA results from the tests of Group 3 for the column (a) and beam (b) datasets when $S_a(T_1)$ is the IM.

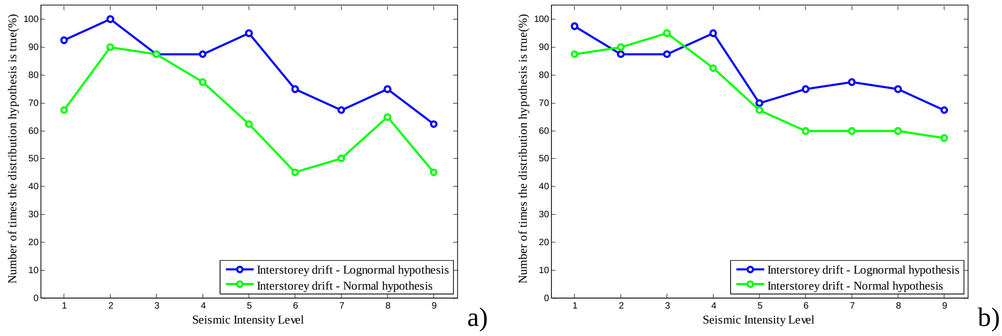

Figure 4. APA results from the tests of Group 3 for the Δ datasets when PGA (a) and $S_a(T_1)$ (b) are the IMs.

In order to examine the reasons behind some of the lower APA results previously referred, some example situations exhibiting less satisfactory GOF results are discussed next. It is noted that an extensive structure-by-structure presentation of the analysis of all the demand parameters represents a prohibitive amount of information to be shown herein. Hence, for the sake of brevity, only a few selected cases are referred. To illustrate a situation where GOF results for φ are less satisfactory, Fig. 5 a) presents the APA results obtained from the application of the tests from Groups 1, 2 and 3 for the columns of the REG10 structure when PGA is the selected IM. As can be observed, the GOF results of the Group 3 tests for the lognormal distribution hypothesis do not meet the 75% APA threshold for some of the intensities and, for the lower intensities, the normal distribution hypothesis yields better APA results. Moreover, it can also be observed that, particularly for intensities 4 and 5, the results from the tests of Group 1 and 2 are considerably different. Since the APA results from the Group 2 tests are higher, such differences indicate that the demand datasets are asymmetric due to the existence of outliers. In order to reduce the influence of the referred outlying observations and improve the APA results, several data processing measures (DPMs) were defined. Since the thorough analysis of each individual dataset is beyond the scope of the present study, the selected measures are global data processing approaches to be applied to all the datasets of a given demand parameter and for a certain intensity level. Based on the observation of some of the individual demand datasets of several intensity levels, the following three global DPMs were considered:

- Data Processing Measure 1 (DPM 1) – Exclusion of the three lowest values from a given dataset (in absolute values, if the demand parameter is negative)
- Data Processing Measure 2 (DPM 2) – Exclusion of the three largest values from a given dataset (in absolute values, if the demand parameter is negative)
- Data Processing Measure 3 (DPM 3) – Exclusion of the three largest and of the three lowest values from a given dataset (in absolute values, if the demand parameter is negative)

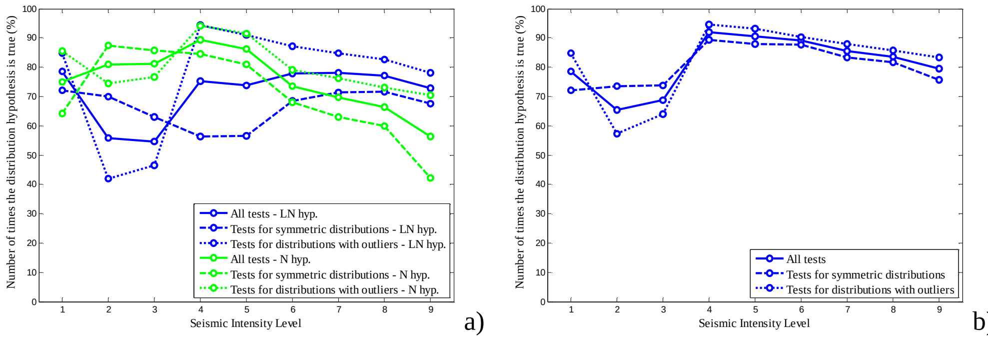

Figure 5. APA results for the REG10 column φ when PGA is the IM (LN hyp. is the lognormal hypothesis and N hyp. is the normal hypothesis) (a) and enhanced APA results for the lognormal hypothesis only (b).

For the referred case of structure REG10, the DPMs were applied to the column φ datasets of all the IM levels to obtain the highest possible APA results for the lognormal distribution hypothesis since, according to the preliminary conclusions previously referred, this distribution is favoured for the probabilistic modelling of φ. For intensity 1, the original APA values from Fig. 5 a) are the highest, for intensities 2 to 5, the highest APA results were obtained by applying the DPM 1, for intensity 6, the highest values were obtained by applying the DPM 3, for intensities 7 and 8, the highest values were obtained by applying the DPM 2, and for intensity 9, the highest values were obtained by applying the DPM 2 to the positive φ datasets and the DPM 3 to the negative φ datasets. To illustrate these findings, Fig. 5 b) presents the enhanced APA results of Fig. 5 a) only for the lognormal distribution hypothesis. As can be observed, the APA results are now generally higher. Nonetheless, the results of intensities 2 and 3 are still below the 75% threshold. For these intensities, a number of datasets has been found to be mostly symmetric, since the results of the Group 1 tests are higher than those of Group 2, while other datasets are negatively skewed and exhibit more than 3 outliers (at each end of the datasets or at one end only), meaning that the application of the DPM 1 may be insufficient to lead to APA results that meet the target threshold. For the remaining intensities, the Group 1 and 2 test results are closer, meaning that the censored datasets are more symmetric and less influenced by outliers.

To illustrate another problematic situation, a case of the V demand is addressed. Generally, the APA values are considerably lower for V than for the other demand parameters and, as previously observed, both distribution hypotheses yield similar results. The reason behind the lower APA values obtained for V is directly connected to the expected evolution of its values. Since the post-yield stiffness of a structural member is usually low, the spread of the V demand distribution tends to be very small when a given structural member has yielded at both ends. In such cases, two conditions were found to occur. In the first condition, some sections exhibited a V distribution which was found to be very irregular and, in some cases, almost uniform. In the second condition, some sections exhibited a V demand distribution with a set of values following the proposed distribution hypotheses mixed with a considerable number of outliers. This second condition was observed, for example, in sections where, for a particular IM level, some of the considered ground motions led to yielding while others did not. To illustrate a situation where both conditions can be observed, Fig. 6 presents the V demand for beam sections of the REG10 structure for intensity 9, when $S_a(T_1)$ is the IM. In this case, most of the left and right beams of the lower storeys have yielded, thus leading to the low dispersion of the demand, while the second condition occurs in several central beams, namely in the top storeys. To further observe the influence of these two conditions, the APA results for that case are presented in Fig. 7 a) where a clear decrease of the APA values can be observed as the intensity level increases, i.e. as the nonlinearity extends to more beams. Moreover, the application of the DPMs does not lead to enhanced APA results that are much higher, as can be seen from the results presented in Fig. 7 b).

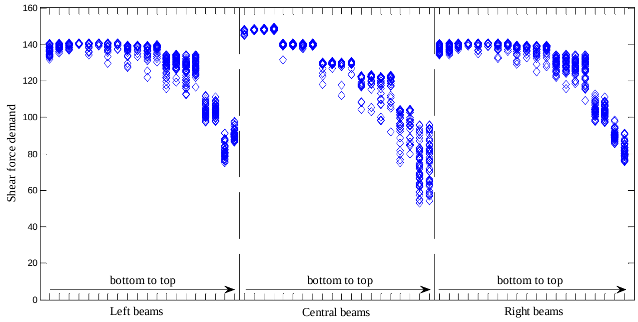

Figure 6. Individual V datasets for the REG10 beam sections structure for intensity 9, when $S_a(T_1)$ is the IM

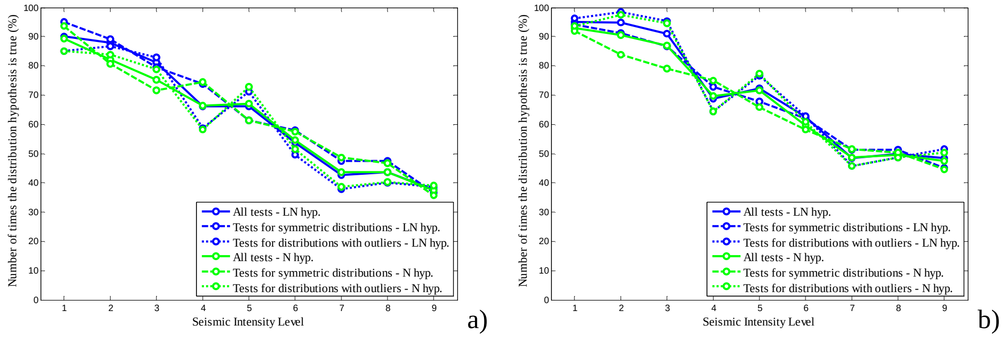

Figure 7. V (a) APA results and (b) enhanced APA results for the beams of REG10 when $S_a(T_1)$ is the IM (LN hyp. is the lognormal hypothesis and N hyp. is the normal hypothesis)

Based on these results, the normal and the lognormal distribution appear to be inadequate to model the V demand distribution in some cases. Still, the effect of considering them in such cases is addressed in the following to determine if their use can be foreseen. This analysis is based on the comparison of the V fragility values λ of selected sections obtained from the empirical and a fitted cumulative distribution function (CDF) of the demand, F_{DP}. The λ values were obtained by Eq. (3.3.1), where f_c is the probability density function (PDF) of the capacity, for higher IM levels with lower APA results.

$$\lambda = \int_0^{\infty} \left(1 - F_{DP}(\alpha)\right) f_C(\alpha)\, d\alpha \tag{3.1}$$

To illustrate the results found, a few examples are presented herein for the normal distribution. Similar results were found for the lognormal distribution. To be able to compare λ values of different ranges, 2 capacity PDFs were considered for each section. The PDFs were defined by a normal distribution with a mean value C_μ and a standard deviation compatible with a coefficient of variation of 5%. For the selected sections (one column section of the ICONS frame for intensity 8 and one beam section of the IRREG10 frame for intensity 9, both when $S_a(T_1)$ is the IM), Table 3.1 presents the λ values obtained for the empirical (λ_{emp}) and the fitted normal (λ_{fit}) demand CDFs, considering the 2 referred capacity PDFs. As can be seen, the λ_{emp} values are always larger, i.e. on the safe side. The relative errors ε between λ_{emp} and λ_{fit} are also presented in Table 3.1. As can be seen, the ε values of the ICONS and IRREG10 sections are similar for both capacity PDFs. From the analysis, it was concluded that although the normal (and the lognormal) distribution might not lead to adequate fits to the demand or to adequate APA results, the errors of considering this distribution are acceptable and on the safe side.

Table 3.1. λ values of the selected sections with the considered values of Cμ and the relative errors ε.

	ICONS column section	IRREG10 beam section
Empirical CDF	λ_{emp} = 3.14E-3 (C_μ = 29kN) λ_{emp} = 1.63E-2 (C_μ = 28kN)	λ_{emp} = 9.75E-3 (C_μ = 155kN) λ_{emp} = 3.58E-2 (C_μ = 150kN)
Normal fitted CDF	λ_{fit} = 3.56E-3; ε = *13.4%* (C_μ = 29kN) λ_{fit} = 1.86E-2; ε = *14.1%* (C_μ = 28kN)	λ_{fit} = 1.16E-2; ε = *19.0%* (C_μ = 155kN) λ_{fit} = 4.47E-2; ε = *24.9%* (C_μ = 150kN)

To emphasize the influence of the DPMs on the APA values, GOF results showing the enhanced APA values of Figs. 3 a) and b) and Figs. 4 a) and b) are shown in Figs. 8 a) and b) and Fig. 9, respectively. Based on the previous findings, only the lognormal hypothesis is considered for the φ, θ and Δ. The enhanced APA results indicate that for the φ, θ, and V, the best results are obtained when $S_a(T_1)$ is the IM. For Δ such trend is not as clear, though $S_a(T_1)$ is better for more intensities. It is also seen that only the V demand APA results do not meet the 75% threshold value, namely for most intensities in beams and for intensity 4 of the columns. Still, the lognormal and normal distributions are accepted based on the fragility analysis previously referred. Finally, the enhanced APA results emphasize the influence of the outliers, which implies that the distribution parameters should be obtained by robust methods.

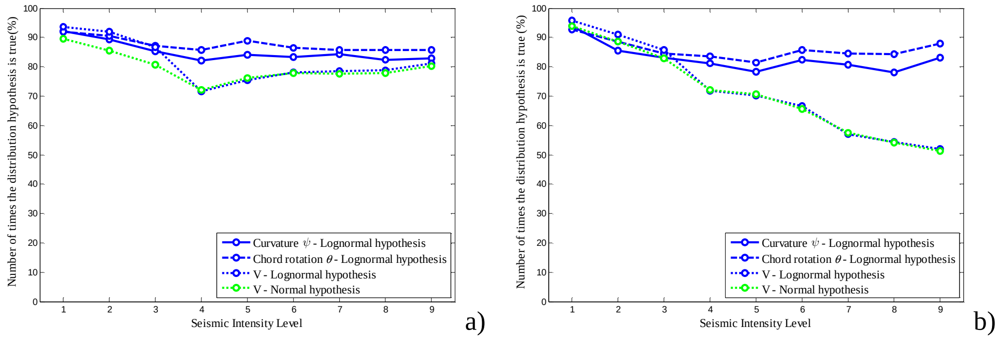

Figure 8. Group 3 tests enhanced APA results for the column (a) and beam (b) datasets when $S_a(T_1)$ is the IM.

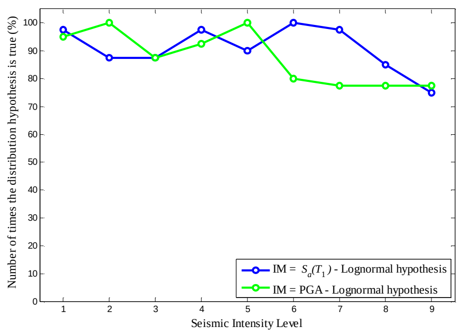

Figure 9. Enhanced APA results from the tests of Group 3 for the Δ datasets for both IMs

5. CONCLUSIONS

The hypothesis that a lognormal or a normal distribution could adequately model the probabilistic distribution of several seismic demand parameters was evaluated. The selected demand parameters were the φ, the θ, the V and the Δ. Several structures were analysed for suites of 50 ground motions. The chosen records were scaled for several intensities to evaluate the referred hypotheses for different hazard levels, and two different IMs were also considered to evaluate the influence of this parameter.
The GOF results were obtained using appropriate statistical methods and were presented in terms of APA values data for the considered levels of seismic intensity and for both distribution hypotheses. The APA results indicated that the lognormal distribution is suitable for the probabilistic modelling of the φ, the θ and the Δ demands. With respect to the V demand, the results indicated that both

distributions have the same potential to model its probabilistic distribution. Since the V APA results were not totally satisfactory, a fragility analysis was performed to determine if the consideration of these distributions would lead to unacceptable errors. From this analysis it was concluded that the selected distributions hypotheses lead to fragility values that are on the safe side with acceptable errors. With respect to the type of IM, $S_a(T_1)$ was seen to be generally more adequate than PGA for the purpose of obtaining demand distributions more compatible with the selected distribution hypotheses.

Finally, the analysis of the individual datasets indicated that outliers occur in many cases. Enhanced APA results were obtained after applying several DPMs to the datasets. The differences between the original and the enhanced APA results emphasized the influence of the outliers, thus implying that robust methods should be used to determine the distribution parameters and minimize their effects.

AKCNOWLEDGEMENT

Financial support of the Portuguese Foundation for Science and Technology, namely through the PhD grant of the first author (SFRH/BD/32820/2007), is gratefully acknowledged.

REFERENCES

Carvalho, A., Zonno, G., Franceschina, G., Bile Serra, J., Campos Costa, A. (2008). Earthquake shaking scenarios for the metropolitan area of Lisbon. *Soil Dynamics and Earthquake Engineering* **28:5**, 347-364.

Arêde, A. and Pinto, A.V. (1996). Reinforced concrete global section modelling: definition of skeleton curves. Special Publication No.I.96.36. JRC, Ispra, Italy.

Athanassiadou, C.J. (2008). Seismic performance of R/C plane frames irregular in elevation. *Engineering Structures* **30:5**, 1250-1261.

Bonett, D.G. and Seier, E. (2002). A test of normality with high uniform power. *Computational Statistics & Data Analysis* **40:3**, 435-445.

Brys, G., Hubert, M and Struyf, A. (2008). Goodness-of-fit tests based on a robust measure of skewness. *Computational Statistics* **23:3**, 429-442.

Carvalho, E.C., Coelho, E., Campos-Costa, A. (1999). Preparation of the full-scale tests on reinforced concrete frames - Characteristics of the test specimens, materials and testing conditions. ICONS report, Innovative Seismic Design Concepts for New and Existing Structures, European TMR Network, LNEC, Lisbon.

CEB (1996). RC frames under earthquake loading. Bulletin no.231, Comité Euro-International du Béton.

Chen, L. and Shapiro, S.S. (1995) An alternative test for normality based on normalized spacings. *Journal of Statistical Computation and Simulation* **53:3**, 269-287.

Coin, D. (2008) A goodness-of-fit test for normality based on polynomial regression. *Computational Statistics & Data Analysis* **52:4**, 2185-2198.

Cornell, C.A., Jalayer, F., Hamburger R.O., Foutch, D. (2002). The probabilistic basis for the 2000 SAC/FEMA steel moment frame guidelines. *Journal of Structural Engineering* **128:4**, 526-533.

Elamir, E.A.H. and Seheult, A.H. (2003). Trimmed L-moments. *Computational Statistics & Data Analysis* **43:3**, 299-314.

EC8-1 (2009). NP ENV 1998-1. Portuguese National Annex to Eurocode 8: Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings. Instituto Português da Qualidade.

Ferracuti, B., Pinho, R., Savoia, M. and Francia, R. (2009). Verification of displacement-based adaptive pushover through multi-ground motion incremental dynamic analyses. *Engineering Structures* **31:8**, 1789-1799.

Gel, Y.R., Miao, W. and Gastwirth, J.L. (2007) Robust directed tests of normality against heavy-tailed alternatives. *Computational Statistics & Data Analysis* **51:5**, 2734-2746.

Goda, K., Hong, H.P., Lee, C.S. (2009). Probabilistic characteristics of seismic ductility demand of SDOF systems with Bouc-Wen hysteretic behaviour. *Journal of Earthquake Engineering* **13:5**, 600-622.

NGA (2009). Pacific Earthquake Engineering Research Center: NGA Database. http://peer.berkeley.edu/nga/.

Romão, X., Delgado, R., Costa, A. (2010). An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation* **80:5**, 545-591.

Ruiz-Garcia, J., Miranda, E. (2010). Probabilistic estimation of residual drift demands for seismic assessment of multi-story framed buildings. *Engineering Structures* **32:1**, 11-20.

Shome N., Cornell C.A. (1999). Probabilistic seismic demand analysis of nonlinear structures. Report RMS-35, Reliability of Marine Structures Program. Stanford University. Stanford, California.

Varum, H. (1997). Numerical model for the seismic analysis of reinforced concrete plane frames. MSc Thesis, Fac. of Eng. of the Univ. of Porto, Porto, Portugal (in Portuguese).